



**QUEEN'S  
UNIVERSITY  
BELFAST**

## **Spatial requirement for phenylacetone monooxygenases for transformation of non-native linear substrates**

Carvalho, A., Dourado, D., Skvortsov, T., Miguel, D. A., Ferguson, L., Quinn, D., Moody, T., & Huang, M. (2018). Spatial requirement for phenylacetone monooxygenases for transformation of non-native linear substrates. *Physical Chemistry Chemical Physics*, 1-32. <https://doi.org/10.1039/C7CP07172H>

**Published in:**  
Physical Chemistry Chemical Physics

**Document Version:**  
Peer reviewed version

**Queen's University Belfast - Research Portal:**  
[Link to publication record in Queen's University Belfast Research Portal](#)

**Publisher rights**  
Copyright 2017 Royal Society of Chemistry. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

**General rights**  
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**  
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

# Spatial requirement for phenylacetone monooxygenases for transformation of non-native linear substrates

*Alexandra T.P. Carvalho,<sup>a,b, ξ‡</sup> Daniel F.A.R. Dourado,<sup>a,b‡</sup> Timofey Skvortsov,<sup>b,c</sup> Miguel de Abreu,<sup>b</sup> Lyndsey J. Ferguson,<sup>b</sup> Derek J. Quinn,<sup>b</sup> Thomas S. Moody,<sup>b</sup> Meilan Huang<sup>a,\*</sup>*

a. School of Chemistry and Chemical Engineering, Queen's University, David Keir Building,  
Stranmillis Road, Belfast BT9 5AG, Northern Ireland, UK

b. Almac Sciences, Department of Biocatalysis and Isotope Chemistry, Almac House, 20 Seagoe  
Industrial Estate, Craigavon BT63 5QD, Northern Ireland, UK

c. School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97  
Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK

ξ Current address: CNC—Center for Neuroscience and Cell Biology, University of Coimbra,  
3004-504 Coimbra, Portugal

## ABSTRACT

Phenylacetone monooxygenase is the most stable and thermo-tolerant member of the Baeyer-Villiger monooxygenases family, and therefore it is an ideal candidate for the synthesis of industrially relevant ester or lactone compounds. However, its limited substrate scope has largely limited its industrial applications. Linear substrates are interesting from an industrial point of view, it is thus necessary to identify the essential spatial requirement for achieving high conversions for non-native linear substrates. Here using molecular dynamics simulations, we compared the conversion of a non-native linear substrate 2-octanone and the native substrate phenylacetone, catalyzed by the WT enzyme and a quadruple variant P253F/G254A/R258M/L443F that exhibits significantly improved activity towards 2-octanone. We uncovered that a remarkable movement of L289 is crucial for a reshaping of the active site of the quadruple variant so as to prevent the aliphatic substrate from moving away from the C4a-peroxyflavin, thus enabling it to keep a catalytically relevant pose during oxygenation process. By performing steady-state kinetic analysis of two single-mutation variants at position 258, we further validated that the L289 reposition is attributed to the combined effect of quadruple mutations. In order to further explore the substrate scope of PAMO we also studied the binding of cyclopentanone and 2-phenylcyclohexanone, which are the typical substrates of CPMO in group I and CHMO in group III, respectively. Our study provides fundamental atomic-level insights in rational engineering of PAMO for wide applications in industrial biocatalysis, in particular, in the biotransformation of long-chain aliphatic oils into potential biodiesels.

**KEYWORDS:** bioengineer, phenylacetone monooxygenases, MD simulations, steady-state kinetic analysis, single-point mutation.

## 1. Introduction

Baeyer-Villiger monooxygenases (BMVOs) are industrially important enzymes that are able to oxidize a variety of substrates such as cyclic or acyclic ketones into the corresponding lactones or esters.<sup>1,2,3</sup> Based upon their amino acid sequence and their selectivity for substrates, BVMO's have been classed into five major groups.<sup>4</sup> The prototype of the BVMO enzyme family is the *Thermobifida fusca* phenylacetone monooxygenase (PAMO), a group II BVMO.<sup>5</sup> This enzyme is particularly thermo-stable<sup>6</sup> and tolerant for different solvents.<sup>5</sup> However, it has very limited substrate scope compared with other BVMOs such as CPMO in group I or CHMO in group III, making it an ideal target for protein engineering.

Attempts to introduce mutations on PAMO by employing saturation mutagenesis approaches were made previously aiming to increase its substrate scope. Except for a directed evolution study where the effect of distal residues was analyzed,<sup>7</sup> most of the mutagenesis studies were focused on the active site residues including the loop residues 440-446.<sup>8,9</sup> The enantioselectivities for a range of ketone, amines and sulfides substrates were improved by mutating the loop residue M446 into a glycine.<sup>10</sup> Increased activity towards bulky substrate 2-phenylcyclohexanone and analogues were achieved by deleting the loop residues S441 and A442 and randomly mutating positions 441 to 444, although the enantioselectivity was low.<sup>11</sup> Mutants with higher *E*-values for 2-phenylcyclohexanone were obtained by randomization of position 440, which can also convert bulky substituted cyclohexanones.<sup>6</sup> Iterative saturation mutagenesis (ISM) studies on five active site residues generated additional variants that were able to convert para-substituted cyclohexanone derivatives. Furthermore, a triple mutant resulted from combination of the single mutants F440P, Q93N and P94D,<sup>7,12</sup> exhibited higher conversion towards cyclohexanone with a minimal trade-off in enzyme thermo-stability. Some of these

mutations severely decreased the conversion of the native substrate.<sup>7</sup> More recently, the screening of a random library composed by 1,500 mutants allowed to identify a PAMO mutant with increased activity towards linear substrates.<sup>13</sup> The best mutant, P253F/G254A/R258M/L443F, was shown to be able to convert several substrates that could not be converted by the WT PAMO. Particularly, it showed a significant increase in activity for a linear aliphatic substrate 2-octanone (19.2-fold increase in  $k_{\text{cat}}/K_{\text{M}}$  in relation to the WT enzyme). The reaction mechanism of BVMO, and particularly PAMO, with native substrate phenylacetone proceeds via the formation of a Criegee intermediate with anionic character, which is subsequently rearranged via the migration of alkyl group to yield the product ester.<sup>14–16</sup> We recently provided an atomic-level understanding of the catalytic mechanism of WT PAMO-catalyzed transformation of the native substrate phenylacetone and the linear substrate 2-octanone, using theoretical QM/MM calculations and Molecular Dynamics simulations.<sup>17</sup> This provides the basis for further investigate the spatial requirement that is essential to accommodate the non-native linear substrate,

Ricinoleic acid is the precursor of castor oil, a main resource in biodiesel production. Recently, a large-scale biotransformation of ricinoleic acid into ester catalyzed by engineered BVMO was reported, which demonstrates the industrial application of BVMO-based whole-cell biocatalysis.<sup>16</sup> Thus it is important to understand the binding of non-native linear substrates to PAMO in order to explore the potential application of engineered BVMO in the catalytic conversion of unsaturated linear substrates in production of biofuels.

Hereby we compared the binding modes of the non-native linear substrate 2-octanone and the native substrate phenylacetone in the WT PAMO and its mutants, using molecular dynamics simulations and steady-state kinetic analysis. We found that a movement of the active site

residue L289 in the mutant P253F/G254A/R258M/L443F is crucial for the reshaping of the catalytic site, enabling the aliphatic substrate to retain a relevant catalytic pose during oxygenation reaction process. The L289 movement is absent in the less active single-mutation variants R258M and R258A. These conformational changes account for the spatial requirement for achieving high conversion for linear substrate and therefore provide important insights for the rational engineering of PAMO to expand its substrate specificity.

## **2. Materials and Methods**

### **2.1 Computational methods**

#### **2.1.1 Modelling**

The initial model system was based on the crystal structure of the flavin-peroxide intermediate of *T. fusca* PAMO (pdb code: 2YLT),<sup>18</sup> which contains the FAD cofactor and the NADP<sup>+</sup> cofactor (the 2-(n-morpholino)-ethanesulfonic acid inhibitor was removed). The C4a-peroxyflavin was modeled based on the crystal pose of the FAD cofactor and in the anionic form was adopted due to the stabilization by the surrounding catalytic residue Arg337 and anionic feature of the subsequently generated Criegee intermediate. The C4a-peroxyflavin, substrate, NADP<sup>+</sup> and Criegee intermediates geometries were optimized using the Gaussian09 program with the B3LYP exchange-correlation functional and the 6-31G(d) basis set in the condensed phase. Point charges were calculated resorting to the RESP method<sup>19</sup> from HF/6-31G(d) single point energy calculations.

#### **2.1.2 Molecular Docking**

Molecular docking was performed using the AutoDock 4.2 suite of programs with the

Lamarckian genetic algorithm (LGA).<sup>20</sup> A grid box was centered on the oxygen of the C4a-peroxy group. R337 was set to be flexible. A total of 100 LGA runs were carried out for each ligand:protein complex. The population was 300, the maximum number of generations was 27,000 and the maximum number of energy evaluations was 2,500,000.

### **2.1.3 Molecular dynamics**

MD simulations were performed using the Amber molecular dynamics program (AMBER14)<sup>21</sup> with the parm99SB<sup>22</sup> and GAFF<sup>23</sup> force fields. The structures were placed within an octahedral box of TIP3P<sup>24,25</sup> waters and counter ions were added to make the entire system neutral. The systems were subjected to two initial energy minimizations and to 500 ps of equilibration in a NVT ensemble using Langevin dynamics with small restraints on the protein (10 kcal/mol) to heat the system from 0K to 300 K. Production simulations were carried out at 300 K in the NPT ensemble using Langevin dynamics with a collision frequency of 1.0 ps<sup>-1</sup>. Constant pressure periodic boundary conditions were imposed with an average pressure of 1 atm. Isotropic position scaling was used to maintain pressure with a relaxation time of 2 ps. The time step was set to 2 fs. SHAKE constraints were applied to all bonds involving hydrogen atoms.<sup>26,27</sup> The particle mesh Ewald (PME) method<sup>28</sup> was used to calculate electrostatic interactions with a cutoff distance of 10 Å. For each reactant complex 5 simulations with random initial velocities were run.

## **2.2 Experimental Methods**

### **2.2.1 Chemicals, Strain and Plasmid**

Chemicals were purchased from Sigma Aldrich UK. The PAMO sequence of 1,638 bp (accession code: Q47PU3) was cloned into the pET-28a(+) plasmid utilising the NdeI and XhoI restriction sites.

### **2.2.2 Site-directed mutagenesis**

Site-directed mutagenesis was performed using the protocol described by Xia *et al.*<sup>29</sup> with the following 5'-overhang oligonucleotides: for the R258A mutation were For-R258A 5'-ACCCAC**G**CGTATCAAGGTCCGAAGAG-3' and Rev-R258A 5'-TTGATAC**G**CGTGGGTGCCACCCGGGGTG-3' and for the R258M were For-R258M 5'-ACCCAC**A**TGTATCAAGGTCCGAAGAG-3' and Rev-R258M 5'-TTGATAC**A**TGTGGGTGCCACCCGGGGTG-3'. After DpnI digestion, 5 µl of the product from each preparation was transformed into the *E. coli* competent cells. The PCR products were also checked with agarose gel electrophoresis. Specific base changes (bold) were confirmed by sequencing.

### **2.2.3 Protein expression and purification**

*E. coli* cultures were incubated overnight in 5mL LB media supplemented with kanamycin (50 mg/mL) in a shaking incubator at 37°C. Cultures were added to 1 L LB media supplemented with kanamycin and incubated at 37°C to OD<sub>600</sub> = 0.5. Cultures were induced by the addition of IPTG (1 mM final concentration) to the culture media and incubated at 25°C for 16h. After incubation, the cell pellets were collected by centrifugation (5000 rpm at 4°C) and resuspended in 100mM Tris buffer pH 7.4. Cells were lysed by sonication on ice 10 x 10 s separated by a 1 min interval. Cell debris was removed by centrifugation (30 min at 5000rpm) at 4°C. The protein was purified using the HisPur<sup>TM</sup> Cobalt Resin protocol (Thermo Scientific<sup>TM</sup>). Protein concentration was determined using Bradford protocol described by Sigma-Aldrich. Protein



purification was conducted under mild conditions to prevent the loss of the flavin cofactor. The recombinant proteins displayed brightly yellow colour in the solution, indicating the presence of FAD prosthetic group. The concentration of FAD-bound PAMO was determined by measuring the absorbance of the protein at 441 nm. The extinction coefficient for FAD-bound PAMO ( $\epsilon_{441} = 12.4 \text{ mM}^{-1} \text{ cm}^{-1}$ ) was previously obtained by Fraaije *et al.*<sup>5</sup> The protein concentration of the Bradford assay and the concentration of the FAD-bound PAMO were compared and a molar ratio of 1:1 was obtained for the wild-type enzyme and the R258A and R258M variants.

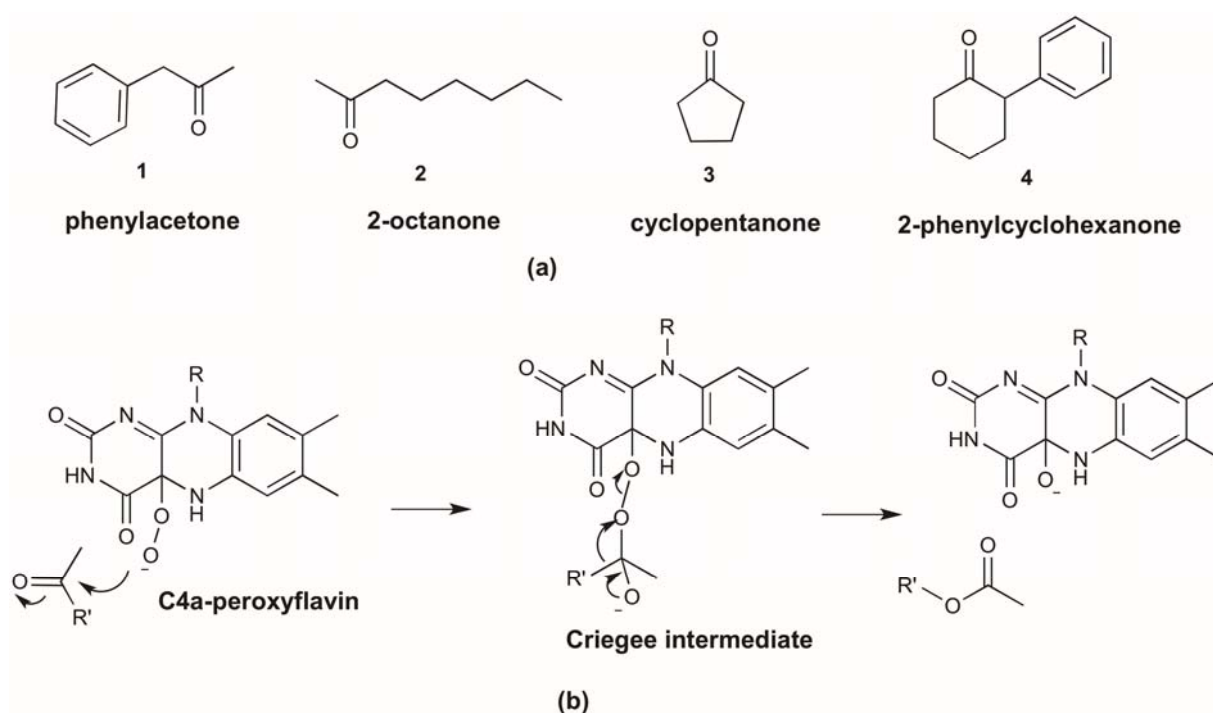
#### **2.2.4 Characterization of enzymes**

The kinetic parameters of the mutants PAMOR258A and PAMOR258M were determined using 96-well plate assay. The conditions used for kinetic analysis assays were as follows: 50 mM Tris·Cl (pH 7.4), 100  $\mu\text{M}$  NADPH, 5% 1,4-dioxane (v/v), 200  $\mu\text{l}$  final volume. All reactions were set up in triplicates and conducted at 25°C. The reactions were initiated by adding NADPH and monitored by continuously measuring the absorbance of the reaction plate samples at 340 nm ( $\epsilon_{340} = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$ ) using Epoch Microplate Spectrophotometer (BioTek, USA). The substrate was in the range of 10 – 5000  $\mu\text{M}$ , while the final concentration of enzymes present in the reaction was in the range 0.1–1.0  $\mu\text{M}$ . The GraphPad Software v6 (GraphPad Software, USA)<sup>30</sup> was used to fit the data and to obtain the kinetic parameters.

### **3. Results and Discussion**

#### **3.1 Binding modes of phenylacetone and 2-octanone in the WT PAMO**

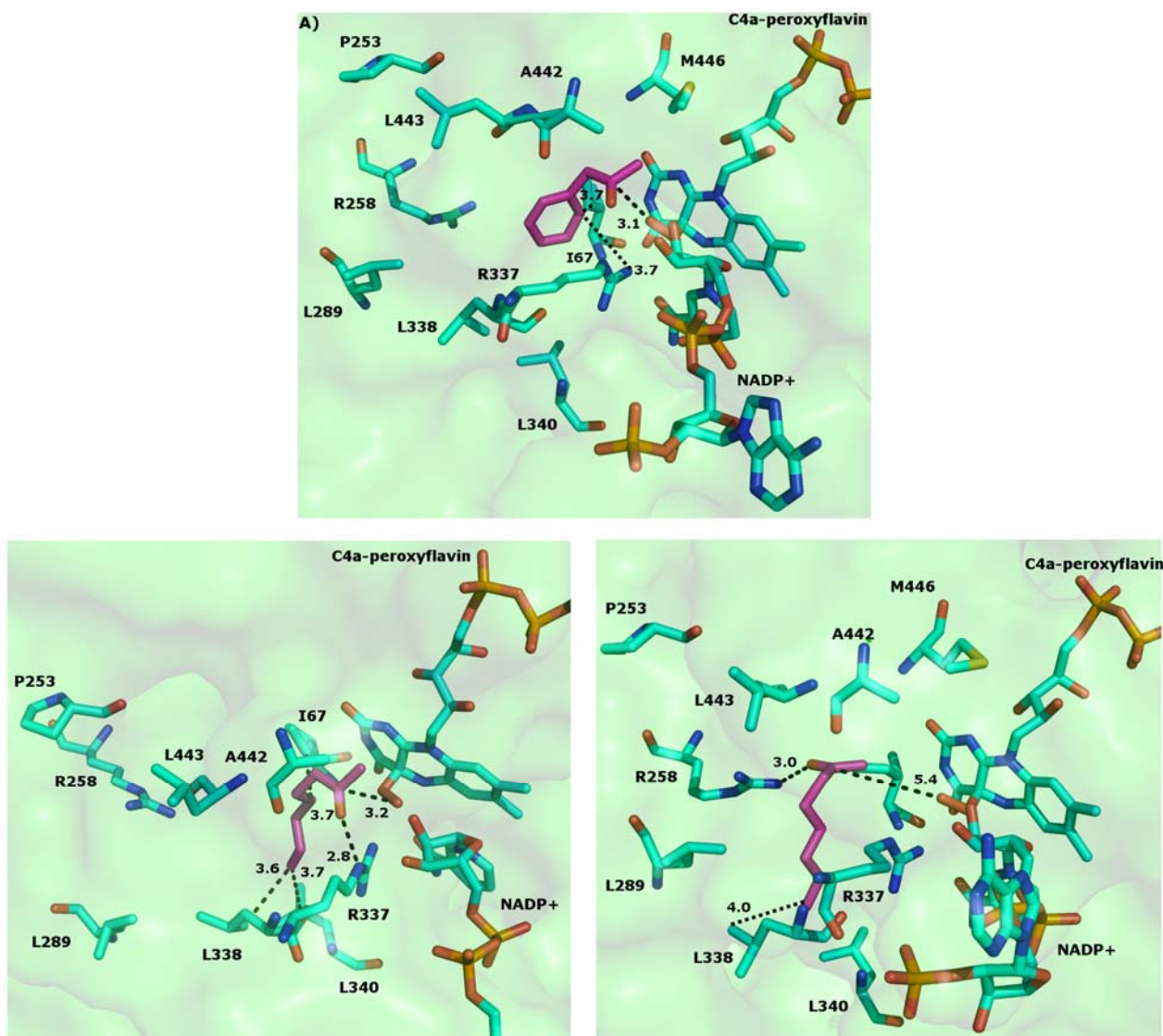
The substrates phenylacetone and 2-octanone share a similar core of atoms that are subjected to nucleophilic attack by the oxygen of the C4a-peroxyflavin FAD cofactor in PAMO (Fig. 1).



**Fig. 1** (a) Substrates studied in this work (b) Substrates, corresponding Criegee intermediates and the product esters in the Baeyer-Villiger oxidation reactions catalyzed by PAMO.

The MD simulations of the WT enzyme bound with the native substrate phenylacetone show that the conserved R337 establishes a cation- $\pi$  interaction with the phenyl ring of the substrate ( $d_{\text{PAC-Phe\_R337-NH2}}(\text{reference structure} - \text{replicas 1-5}) = 3.6 \pm 0.3 \text{ \AA}$ ), and a hydrogen bond with carbonyl oxygen of the substrate ( $d_{\text{PAC-O\_R337-NH}}(\text{reference structure} - \text{replicas 1-5}) = 3.1 \pm 0.2 \text{ \AA}$ ) constituting a binding pose favorable for nucleophilic attack with the carbonyl carbon of the substrate remaining close to the oxygen of the C4a-peroxyflavin cofactor ( $d_{\text{PAC-C2\_C4a-O-O}^-}(\text{reference structure} - \text{replicas 1-5}) = 3.4 \pm 0.2 \text{ \AA}$ ) (Fig. 2A, Fig. S1). It should be noted that in one of the MD replicas (replica 5) the substrate pose changes after 15ns of MD (Fig. S1). Nevertheless, the R337 still establishes a cation- $\pi$  interaction with the substrate ( $d_{\text{PAC-Phe\_R337-NH2}}(20\text{ns average} - \text{replica 5}) = 3.5 \pm 0.8 \text{ \AA}$ ), and carbonyl carbon of the substrate remains close to the oxygen of the C4a-peroxyflavin cofactor ( $d_{\text{PAC-C2\_C4a-O-O}^-}(20\text{ns average} - \text{replica 5}) = 4.0 \pm 1.2 \text{ \AA}$ ).

All these results are in accordance with the experimental data which indicates that mutation of R337 to alanine or lysine abolished the enzymatic activity.<sup>31</sup> The phenyl group of the substrate also interacts with I67 and R337 while the methyl end establishes hydrophobic interactions with A442 and M446, which are located on an active site loop (residues 440-446) and were previously indicated to alter the conversion and enantioselectivity of either bulky cyclohexanone substituents or sulfides.<sup>9,12</sup>



**Fig. 2** MD reference structures of the WT PAMO in complex with A) phenylacetone replica 1 (replica 2, 3, 4 & 5 are similar to replica 1). B) 2-octanone replica 1 (replica 3, 4 & 5 are similar to replica 1). C) 2-octanone replica 2. A

MD reference structure corresponds to the lowest RMSD structure in relation to the average structure of the simulation. Relevant distances are shown (Å).

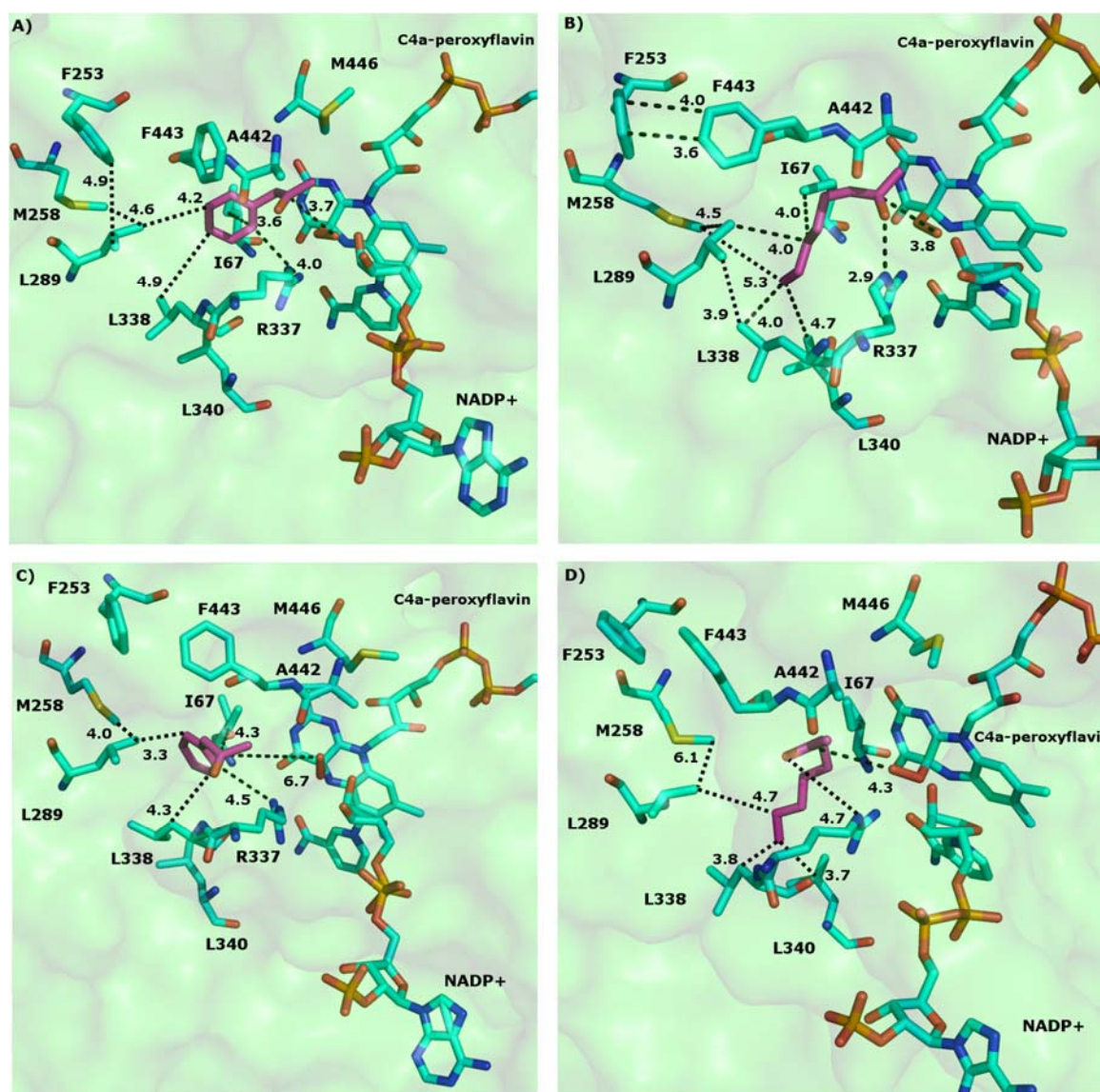
In order to understand the binding of the non-native substrate 2-octanone, the complex structure was also investigated by MD simulations. In four MD replicas 2-octanone binds to the enzyme in similar manner to the native substrate phenylacetone, with the acetone moiety positioned similarly in the productive conformations ( $d_{\text{OCT-C2\_C4a-O-O}^-}$  (reference structure - replicas 1, 3, 4, 5) =  $3.2 \pm 0.1$  Å, Fig. S1 and Fig. 2B). The carbonyl carbon of the substrate is close to the oxygen of the C4a-peroxyflavin cofactor. The methyl group interacts with A442 and M446 of the active site loop while the aliphatic hexyl side chain interacts with I67, L338 and L340 (Fig. 2B). The residue R337 establishes a hydrogen bond with the substrate carbonyl oxygen ( $d_{\text{OCT-O\_R337-NH}}$  (reference structure - replicas 1,3,4,5) =  $2.8 \pm 0.1$  Å). In one of the MD replicas (replica 2) the carbonyl group of 2-octanone moves towards the entrance of the active site pocket and approaches R258, establishing a hydrogen bond with this residue (Fig. 2C) In this pose, the substrate carbonyl group moves away from the sidechain of R337 ( $d_{\text{OCT-O\_R337-NH}}$  (reference structure - replica 2) =  $6.0$  Å). The absence of the favorable strong cation- $\pi$  interaction with R337 observed for the native substrate (Fig. 2A) makes 2-octanone able to move more freely in the large active site pocket and therefore accounts for its lower catalytic efficiency compared with the native substrate phenylacetone. This is in accordance with the previous kinetic study where the  $K_M$  of 2-octanone catalysis is  $\sim 2$  orders of magnitude higher than that of the phenylacetone, which indicates a lower binding affinity.<sup>17</sup>

### **Binding modes of the PAMO variants**

A quadruple mutant P253F/G254A/R258M/L443F was reported to significantly improve enzyme activity for 2-octanone with a  $k_{\text{cat}}/K_M$  of  $9200 \text{ M}^{-1}\text{s}^{-1}$ .<sup>13</sup> In contrast this mutant has a small effect on the conversion of the native substrate phenylacetone since  $k_{\text{cat}}/K_M$  decreases only 1.07-fold in

relation to the WT enzyme.<sup>13</sup> In order to elucidate the effect of the mutations on the catalytic efficiency towards the two distinct substrates, MD simulations were conducted for the quadruple mutant P253F/G254A/R258M/L443F.

MD simulations of the native substrate phenylacetone in the quadruple mutant show that the substrate is nested in a location favorable for the reaction to happen in four of the MD replicas (Fig. 3A), ( $d_{\text{PAC-C2\_C4a-O-O}^-}$ (reference structure - replicas 1,2,,4,5) =  $3.4 \pm 0.3$  Å, Fig. S1). Similar to the binding in the WT enzyme, the phenyl ring of the substrate is orientated towards the entrance of the catalytic site. L289 moves closer to the residues R258M and P253F. As a result, the phenyl ring of the substrate forms favorable hydrophobic interactions with I67, L289 and R337 (Fig. 3A). In one of the MD replicas, the substrate moves away from the FAD peroxy group ( $d_{\text{PAC-C2\_C4a-O-O}^-}$ (reference structure – replica 3) =  $6.7$  Å, Fig. 3C) and from R337, and in the meantime, approaches L289 and L338, resulting in a non-catalytically pose.



**Fig. 3** MD reference structures of the P253F/G254A/R258M/L443F variant in complex with: A) phenylacetone replica 1 (replicas 2, 4 and 5 are similar to replica 1); B) 2-octanone replica 1 (replica 3, 4 and 5 are similar to replica 1); C) phenylacetone replica 3; D) 2-octanone replica 2. A MD reference structure corresponds to the lowest RMSD structure in relation to the average structure of the simulation. Relevant distances are shown (Å).

The MD simulations of the quadruple mutant in complex with 2-octanone showed that in four MD replicas the substrate pose (Fig. 3B & 3D) is similar to what was observed for the WT enzyme (Fig. 2B). The average distance between the oxygen of the C4a-peroxyflavin and

substrate carbonyl carbon, is 3.6Å ( $d_{\text{OCT-C2\_C4a-O-O}^-}$ (reference structure - replicas 1,3,4,5)). However, the substrate in the quadruple mutant establishes more intermolecular interactions within the active center pocket than the WT enzyme, indicating a higher binding affinity. This is in accordance with the increase of  $\sim 1$  order of magnitude in  $K_M$  observed for this mutant.<sup>13,17</sup> L443F is located in the active site loop 440-446 and P253F along with G254A are located in a nearby loop (Fig. S2). The aromatic side chains of P253F and L443F in the mutant form a pi-pi stacking interaction that causes the active site loop shifts slightly away from the C4a-peroxyflavin. Interestingly, this small displacement makes the residue L289 undergo a significant movement towards the entrance of the catalytic site and approach R258M (Fig. 3B, S2). Replacing the loop residue G254 by an alanine residue in the quadruple mutant might increase the rigidity of the loop where it is located, and therefore help to maintain the favorable pi-stacking interaction formed between P253F and L443F. The alkyl tail of the substrate in the variant establishes stronger interactions with the hydrophobic residues I67, L289, L338 and L340 (Fig. 3B). Similar to what was observed in the WT enzyme, the substrate carbonyl oxygen interacts with R337 through a hydrogen bond interaction ( $d_{\text{OCT-O\_R337-NH}}$  (reference structure - replicas 1,,3,4,5) =  $3.4 \pm 1.0$  Å).

It is worth noting that although the linear aliphatic substrate can nest nicely in the newly generated hydrophobic pocket (Fig. 3B), the carbonyl head of the substrate may still deviate from the ideal catalytic pose (Fig. 3D). In one of the MD simulation replicas, the 2-octanone carbonyl oxygen switches to L289 and is involved in a weaker hydrogen bond with R337 sidechain ( $d_{\text{OCT-O\_R337-NH}}$  (reference structure – replica 2) = 4.7 Å). The distance between the oxygen of the C4a-peroxyflavin and substrate carbonyl carbon is 4.3Å, Fig. 3D). Nevertheless, this substrate pose is considerably more catalytically relevant than the one found for the WT enzyme (replica 2, Fig. 2C). Thus it appears that when the linear substrate is bound in PAMO, L289 may adjust



itself between R258M and R337 and a catalytically pose may be achieved by the spatial arrangement of this residue.

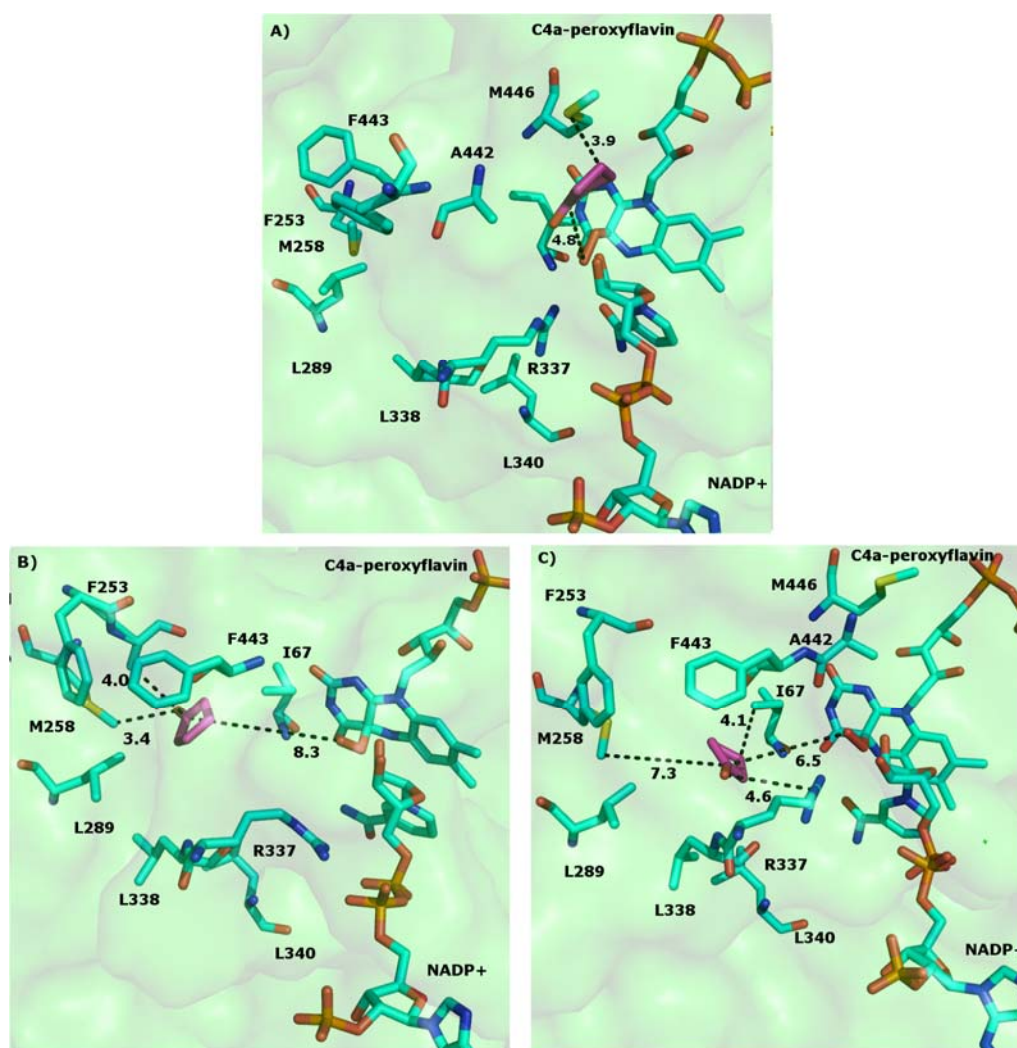
### 3.3 Profiling of non-linear substrates in the quadruple PAMO variant

Compared with other BVMOs, the WT PAMO exhibited limited substrate scope.<sup>13</sup> In order to understand the substrate scope in PAMO, we also studied the binding modes of two substrates, cyclopentanone and 2-phenylhexanone, which are the typical substrates of CPMO in group I and CHMO in group III, respectively.

First, we studied the binding of these substrates in the WT PAMO. For cyclopentanone, the carbonyl group is positioned towards the peroxy group of the C4a-peroxyflavin in only one replica out of the five MD simulations and even for this replica the distance is about 6.9 Å away from the peroxy oxygen (Fig. S3 & S4), indicating the CPMO substrate cyclopentanone cannot position itself favorably in the WT enzyme. This is in accordance with the poor kinetic data observed for cyclopentanone (Table 1).<sup>13</sup> When cyclopentanone is bound with the quadruple variant, consistent improvement in the substrate binding was not observed. In two MD replicas the cyclopentyl ring of the substrate establishes hydrophobic interaction with the side chain of M446, while the carbonyl group becomes closer to the C4a-peroxyflavin oxygen ( $d_{\text{cyclopentanone-C2\_C4a-O-O}^{\cdot-}}(\text{reference structure} - \text{replicas 1,3}) = 4.9 \pm 0.1 \text{ \AA}$ , Fig. 4A). However, it is still not sufficiently close to the C4a-peroxyflavin cofactor for an efficient nucleophilic reaction to happen. Actually in three other replicas of MD simulations (Fig. 4B and 4C), the substrate moves far away from the cofactor ( $d_{\text{cyclopentanone-C2\_C4a-O-O}^{\cdot-}}(\text{reference structure} - \text{replicas 2,4}) = 7.0 \pm 1.3 \text{ \AA}$ , Fig. 4B and  $d_{\text{cyclopentanone-C2\_C4a-O-O}^{\cdot-}}(\text{reference structure} - \text{replica 5}) = 6.5 \text{ \AA}$ , Fig. 4C). Thus the low activity of the WT and the quadruple variant towards cyclopentanone can be mainly attributed to the unavailability of significant catalytic pose due to the free movement of the small substrate in the large binding



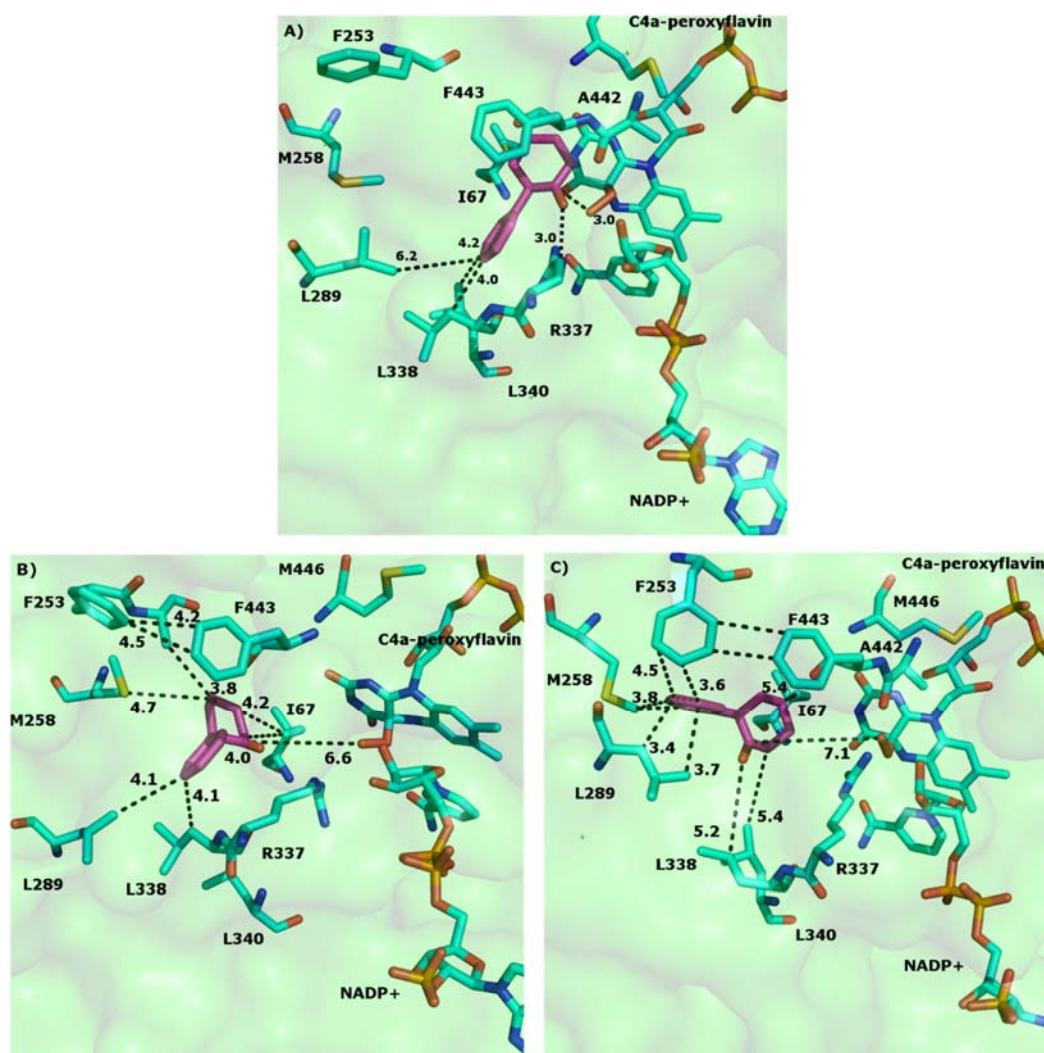
pocket.



**Fig. 4** MD reference structures of the P253F/G254A/R258M/L443F variant in complex with cyclopentanone: A) replica 3 (replica 1 is similar to replica 3); B) replica 4 (replica 2 is similar to replica 4); C) replica 5.

We then examined the binding of 2-phenylhexanone, a typical substrate of CHMO, in PAMO. In the WT PAMO, the phenyl ring of 2-phenylhexanone established favorable interactions with the hydrophobic pocket formed by L289, L338 and L340 while the carbonyl of the substrate is still far away from the cofactor oxygen ( $d_{2\text{-phenylhexanone-C2\_C4a-O-O}^-}$  (reference structure - replicas 1,2,3) =  $5.1 \pm 1.4$  Å,

$d_{\text{2phenylhexanone-C2\_C4a-O-O}^{\sim}(\text{reference structure - replicas 4,5})} = 7.0 \pm 1.0 \text{ \AA}$  (Fig. S3 & S5). In contrast, in one of the MD replicas of the quadruple variant, the carbonyl of the substrate approaches the peroxy group of the cofactor with a favorable distance of 3.0 Å (Fig. 5A); in addition, it forms an H-bond with the guanidino NH of R337 while the hydrophobic interactions with the surrounding hydrophobic residues are maintained. However, in majority of the MD simulated structures (in the other four replicas), it was found that the substrate moves away from the cofactor, albeit establishing new strong hydrophobic intermolecular interactions between the phenyl ring and the mutated residues G254A, R258M and L443F, in addition to the maintained interactions with L289 and L338 ( $d_{\text{2phenylhexanone-C2\_C4a-O-O}^{\sim}(\text{reference structure - replica 2})} = 6.6 \text{ \AA}$ , Fig. 5B and  $d_{\text{2phenylhexanone-C2\_C4a-O-O}^{\sim}(\text{reference structure - replicas 3,4,5})} = 7.2 \pm 0.3 \text{ \AA}$ , Fig. 5C). The tighter binding of the substrate due to the additional hydrophobic interactions with the active centre residues is in line with the observed improvement in the  $K_M$  of the quadruple mutant compared with the WT PAMO (Table 1)<sup>13</sup>. The two non-native substrates 2-octanone and 2-phenylcyclohexanone are converted by the quadruple PAMO variant with reasonably high conversion efficiency. However, the predominant binding poses of the non-linear substrate 2-phenylcyclohexanone are far away from the cofactor, with the phenyl group close to the entrance of the catalytic site. In contrast, the predominant poses of the linear substrate 2-octanone are in proximity of the peroxy oxygen of the cofactor, with the aliphatic chain nested in the hydrophobic pocket formed by I67, L338 and L340, such that the reaction is more liable to happen and associated with higher  $k_{\text{cat}}$  value.

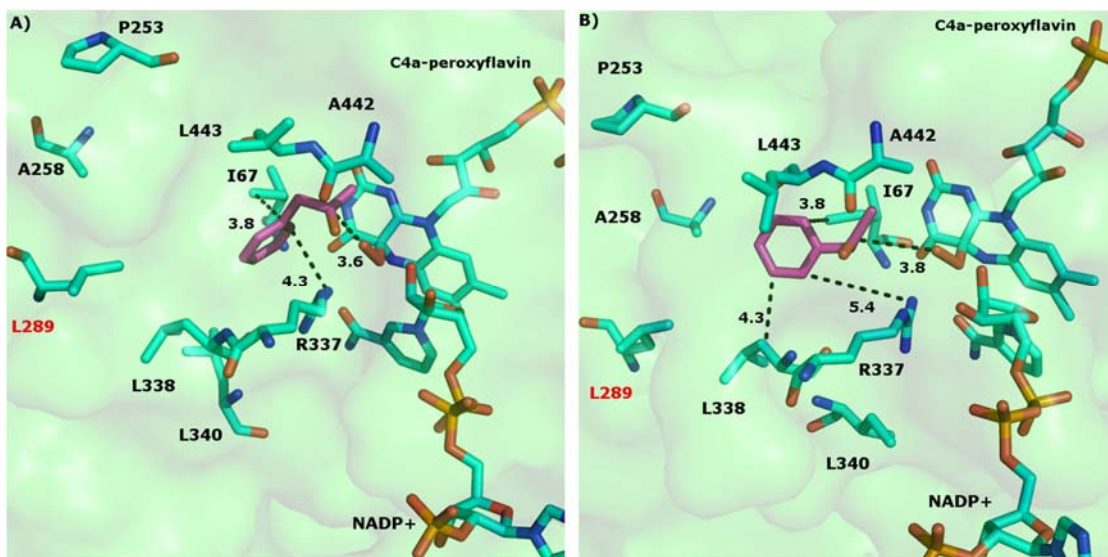


**Fig. 5** MD reference structures of the P253F/G254A/R258M/L443F variant in complex with 2-phenylhexanone: A) replica 1; B) replica 2; C) replica 3 (replica 4 and 5 are similar to replica 3).

### 3.4 Importance of spatial arrangement of L289

In order to demonstrate that the movement of L289 is crucial for 2-octanone to adopt a catalytic pose in the potent quadruple mutant, we conducted MD simulations of two single-mutant variants, R258A and R258M, and also measured the steady-state kinetic parameters. According to a previous mutagenesis study the R258A mutation decreases by 3.2-fold the  $k_{cat}/K_M$  for

phenylacetone<sup>13</sup>. Since there is no structural basis for this *in vitro* data we performed the MD analysis of this mutant.

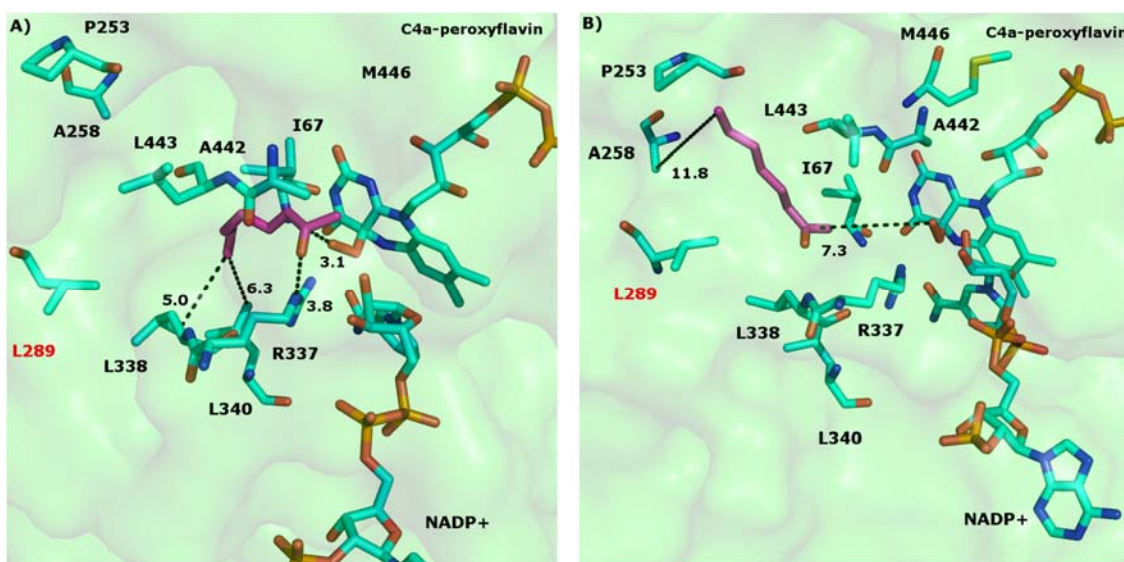


**Fig. 6** MD reference structures of PAMO R258A variant in complex with phenylacetone: A) R258A replica 2 (replicas 3 and 5 are similar to replica 2); B) R258A replica 4 (replica 1 is similar to replica 4). A MD reference structure corresponds to the lowest RMSD structure in relation to the average structure of the simulation. Relevant distances are shown (Å).

Interestingly, we observed that the arrangement of the substrate within the active centre pocket changes in the R258A mutant (Fig 6). The conserved R337 establishes a cation- $\pi$  interaction with the substrate in only three MD replicas ( $d_{\text{PAC-Phe\_R337-NH}_2}$  (reference structure - replicas 2,3,5) =  $3.6 \pm 0.1$  Å) (Fig. 6A). In the other two MD replicas the cation- $\pi$  interaction is lost ( $d_{\text{PAC-Phe\_R337-NH}_2}$  (reference structure - replicas 1,4) =  $4.7 \pm 0.9$  Å), which indicates a decrease in the substrate binding affinity (Fig. 6B). However, in all five MD replicas the carbonyl carbon of the substrate remains close to the oxygen of the C4a-peroxyflavin cofactor ( $d_{\text{PAC-C}_2\text{-C4a-O-O}^-}$  (reference structure - replicas 1,2,3,4,5) =  $3.6 \pm 0.4$  Å, Fig. 6, Fig. S1). This seems to be in accordance with the experimental data, which indicates that the 3.2-fold decrease in the  $k_{\text{cat}}/K_{\text{M}}$  is mostly attributed to an increase in the  $K_{\text{M}}$  (2.8-fold increase).



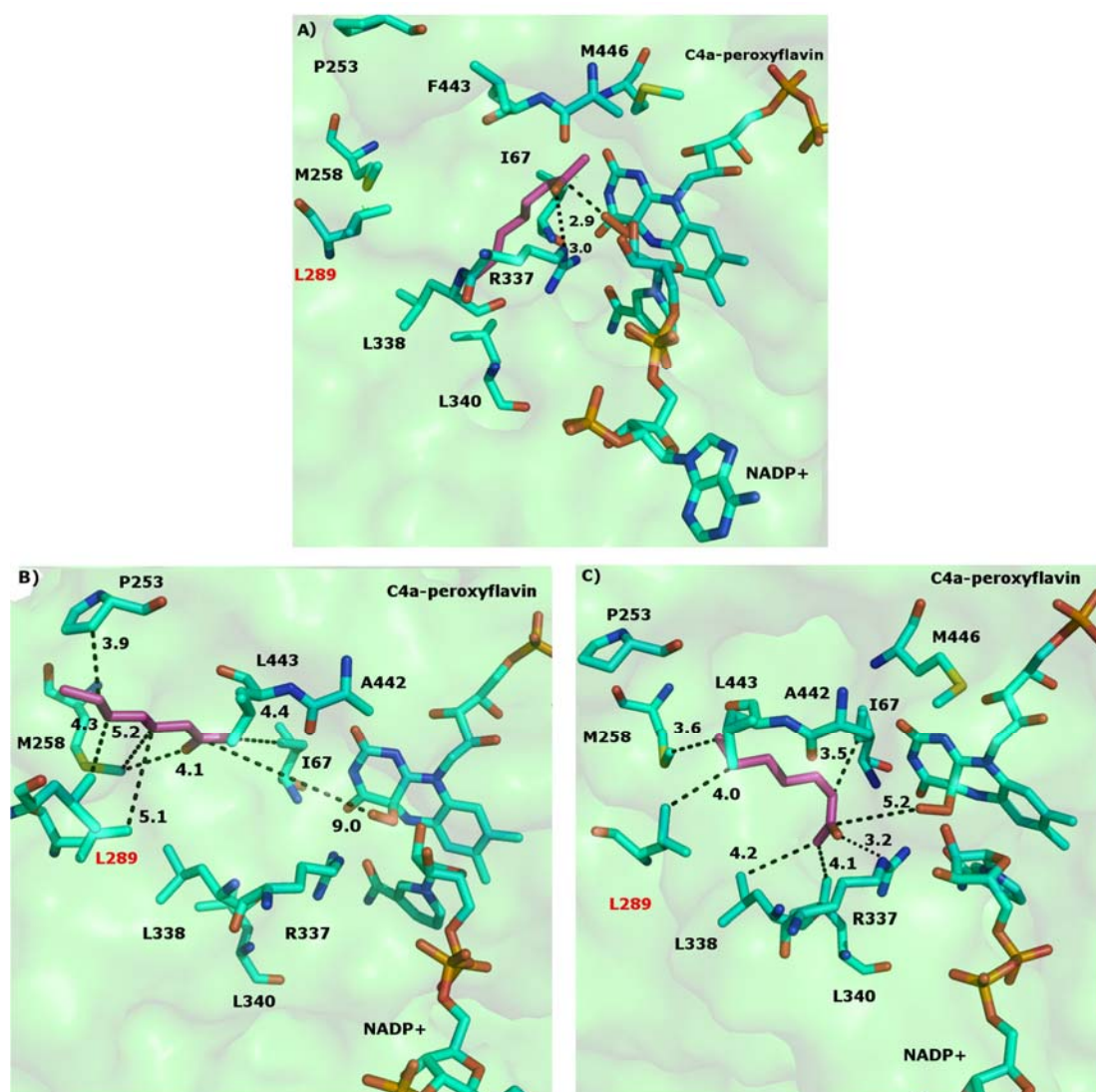
On the other hand, when the substrate 2-octanone binds to the R258A mutant, a significant change in the position of the hexyl tail of the substrate is observed (Fig. 7A & 7B and Fig. S2). It moves away from R337 and L338, and towards the extra space created when the side chain of the R258 was truncated into a methyl group. This change does not affect the distance to the C4a-peroxyflavin oxygen in four of the MD replicas (Fig. 7A). However, in one of the replicas the distance increases to 7.3 Å (Fig. 7B), resulting in a catalytically incompetent pose and may explain the 2.4-fold decrease observed in  $k_{cat}$  compared with the WT enzyme (Table 1).



**Fig. 7** MD reference structures of PAMO R258A variant in complex with 2-octanone: A) R258A replica 1 (replicas 3, 4 and 5 are similar to replica 1); B) R258A replica 2. A MD reference structure corresponds to the lowest RMSD structure in relation to the average structure of the simulation. Relevant distances are shown (Å).

The R258M mutation also significantly affects pose of 2-octanone, since the hexyl tail moves towards M258 (Fig. 8B & 8C and Fig. S2). Moreover, the carbonyl group of 2-octanone moves considerably away from the peroxy-FAD in three of the MD replicas (Fig. 8B & 8C). The average distance between the oxygen of the C4a-peroxyflavin and substrate carbonyl carbon,  $d_{OCT-C2\_C4a-O-O^{\sim}}$  (reference structure - replicas 1,3,5), is  $6.4 \pm 2.2$  Å. This is in accordance with the observed

3.2-fold decrease in  $k_{\text{cat}}$  compared with the WT enzyme (Table 1). On the other hand, this new strong hydrophobic interaction between 2-octanone hexyl tail and the introduced methionine side chain (Fig. 8B & 8C) leads to a tighter binding of the substrate to the active centre pocket, which accounts for the observed 10.7-fold decrease of  $K_M$  when compared with the WT enzyme (Table 1).



**Fig. 8** MD reference structures of PAMO R258M variant in complex with 2-octanone: A) replica 2 (replica 4 is similar to replica 2); B) replica 3; C) replica 1 (replica 5 is similar to replica 1). MD reference structure corresponds to the lowest RMSD structure in relation to the average structure of the simulation. Relevant distances are shown

(Å).

Comparison of the three variants, R258A, R258M and P253F/G254A/R258M/L443F reveals a remarkable movement of L289 for the most potent quadruple variant (Fig. S2), which allows the alkyl tail of the substrate to be held in the hydrophobic region comprised of L289, L338 and L340 such that it cannot move away from the C4a-peroxyflavin cofactor.

**Table 1** Steady-state kinetic parameters of PAMO WT and variants R258A, R258M and P253F/G254A/R258M/L443F for the conversion of 2-octanone. The values obtained are averaged results of three measurements. Kinetic parameters were determined with GraphPad Prism 6; the kinetic curves were fitted using Michaelis-Menten equation for substrate inhibition.

	PAMO	$K_M$ [mM]	$k_{cat}$ [ $s^{-1}$ ]	$k_{cat}/K_M$ [ $M^{-1} s^{-1}$ ]
phenylacetone	WT <sup>13</sup>	0.06	2.4	40
	R258A <sup>13</sup>	0.17	2.1	12
	P253F/G254A/R258M/L443F <sup>13</sup>	1.4	0.04	$3.5 \times 10^4$
2-octanone	WT <sup>17</sup>	3.2	0.22	66
	R258A	2.1	0.091	43
	R258M	0.3	0.067	$2.3 \times 10^2$
	P253F/G254A/R258M/L443F <sup>13</sup>	0.25	2.3	$9.2 \times 10^3$
2-phenylcyclohexanone	WT <sup>13</sup>	>4	>0.07	8
	P253F/G254A/R258M/L443F <sup>13</sup>	0.8	0.3	$3.7 \times 10^2$
cyclopentanone	WT <sup>13</sup>	>1,200	>0.9	0.8
	P253F/G254A/R258M/L443F <sup>13</sup>	1,000	1.6	1.6

Overall, the measured catalytic efficiencies for R258A and R258M are  $43 \text{ M}^{-1} \text{ s}^{-1}$  and  $2.3 \times 10^2 \text{ M}^{-1} \text{ s}^{-1}$ , respectively (Table 1). The catalytic efficiency of R258A,  $43 \text{ M}^{-1} \text{ s}^{-1}$ , is much lower than the one previously measured for the quadruple mutant,  $9.2 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ .<sup>13</sup> The catalytic efficiency of R258M,  $2.3 \times 10^2$ , is one order of magnitude lower than that of the quadruple mutant. For this variant, the favorable improvement on  $K_M$  may have been offset by the decrease in  $k_{\text{cat}}$ .

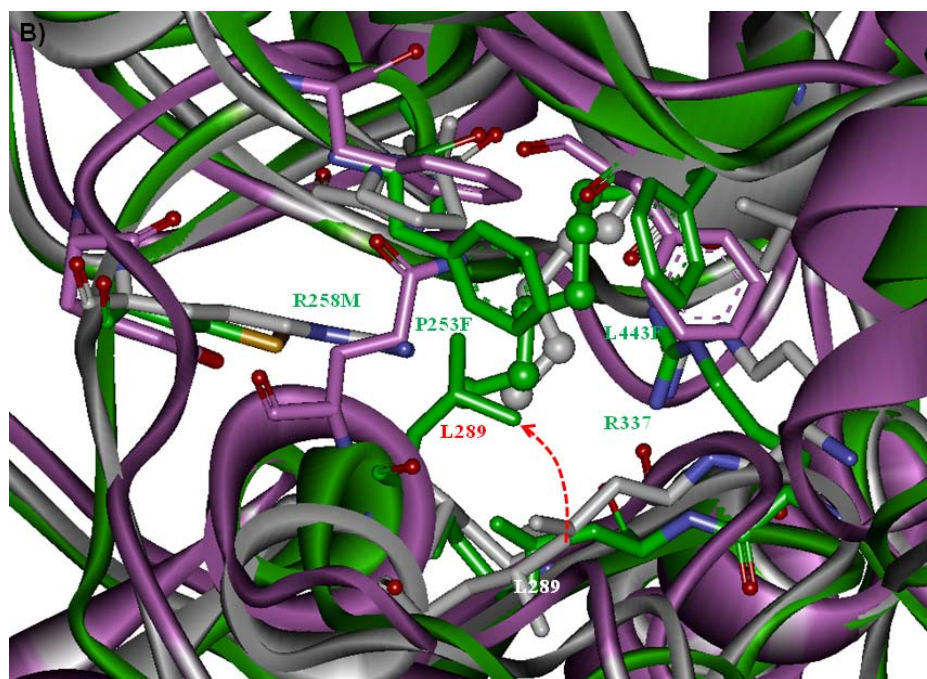
Thus, although these single-point mutations create more space in the active centre, the conformational changes are not sufficient to achieve a high conversion for the linear aliphatic substrate. L289 has to be repositioned so as to create a hydrophobic pocket to limit movement of the substrate, warranting the promising catalytic activity.

It should be noted that despite of the promising activity of the quadruple mutant towards 2-octanone, the carbonyl group of substrate can still move away from the peroxy-FAD, resulting in less catalytically relevant poses (Fig. 3D, Fig. S2). Consequently, further improvement in enzymatic catalytic efficiency may be achieved by restraining the long aliphatic substrates more tightly to the active centre of PAMO.

To validate the facilitating role of L289 in the positioning of the linear substrate in the hydrophobic pocket formed by I67, L338 and L340, L289A was introduced into the quadruple variant and MD simulations were conducted for the P253F/G254A/L289A/R258M/L443F variant. While the aliphatic chain of the linear substrate is positioned in the hydrophobic site (Fig. S6A), the carbonyl group of the substrate is too far away ( $\sim 9.8 \text{ \AA}$ ) such that it is unlikely to be attacked by the peroxy-FAD cofactor. Furthermore, in majority of the MD simulated structures (Fig. S6B, & 5C), instead of interacting with I67, L338 and L340, the aliphatic chain of 2-octanone moves towards the entrance of the catalytic site to establish hydrophobic interaction with M258, resulting in catalytically incompetent poses. Thus L289 plays an essential







**Fig. 9** (A) Sequence alignment of PAMO (PDB code: 2YLT) and *A. flavus* BVMO, BVMO<sub>AFL838</sub> (PDB code: 5J7X). The loop where P253, G254 and R258 are located and the activation loop where 440-446 are underlined by green lines. Three mutation positions are marked by red arrows. (B) Superimposing of the WT (in grey) and quadruple variant of PAMO (in green) with BVMO<sub>AFL838</sub> (in purple).

The crystal structure of BVMO<sub>AFL838</sub> has been resolved (PDB code: 5J7X).<sup>30</sup> Superimposition of the crystal structure of BVMO<sub>AFL838</sub> and the MD reference structure of the quadruple variant (Fig. 9B) disclosed that the phenylalanine residue at position 443 in PAMO aligns itself well with that in BVMO<sub>AFL838</sub>, as well as the phenylalanine residue at position 253. Both P253 and R258 in PAMO are located on a flexible loop near the entrance. Existence of two glycine residues on this loop enables the phenylalanine to adapt itself to form favorable  $\pi$ -stacking interaction with the phenylalanine at the active loop. Thus the common binding mode of the linear substrate 2-octanone in BVMO<sub>AFL838</sub> and the quadruple mutant of PAMO indicate that high conversion efficiency may be achieved by grafting the corresponding non-conserved residue around the substrate from BVMO<sub>AFL838</sub> into PAMO.

#### **4. Conclusions**

The high stability and thermo-tolerance of the flavoprotein PAMO makes it an ideal biocatalyst for the oxidation of ketones. However, its limited substrate scope precludes its broader application in industry. Binding of 2-octanone to an efficient P253F/G254A/R258M/L443F mutant and two single-mutation variants (R258A and R258M) was compared to shed light on further PAMO engineering. Substitution of R258 by alanine or methionine allows the alkyl tail of the substrate to move away from R337 and towards the pocket entrance. Further mutations are required to prevent the substrate to move away from the C4a-peroxyflavin via the reposition of L289. Nevertheless, it should be noted that even though the substrate pose is improved with the aliphatic tail nested in a hydrophobic pocket comprised of L289, L338 and L340, its C1 methyl and carbonyl groups can still move away from the cofactor.

Our study has shown that additional design efforts should be made to improve the binding of the substrate as well as of the corresponding Criegee intermediate. It appears that PAMO can be perfectly adapted to catalyze the oxidation of different ketones with high efficiency as long as the substrates can access the C4-peroxoflavin cofactor and the Criegee intermediates are properly stabilized. In summary, we find that in the PAMO a reshaping of the active site (via the cooperative effect of the mutations and movement of L289) is an essential requirement for improving the binding and conversion of long aliphatic substrates, which may provide significant insight in rationally engineering the enzymes for industrial production of biofuels such as castor oils.

#### **Author Contributions**

‡These authors contributed equally.

### Corresponding Author

\*m.huang@qub.ac.uk

### Acknowledgements

The authors acknowledge the financial support from INVEST NI Research and Development Programme, part financed by the European Regional Development Fund under the Investment for Growth and Jobs programme 2014-2020. We are grateful for the computing resources from QUB high performance computing Centre.

### References

1. ten Brink, G.-J., Arends, I. W. C. E. & Sheldon, R. A. The Baeyer–Villiger Reaction: New Developments toward Greener Procedures. *Chem. Rev.* **104**, 4105–4124 (2004).
2. Schmid, R. D. & Urlacher, V. *Modern Biooxidation: Enzymes, Reactions and Applications*. (John Wiley & Sons, 2007).
3. Torres Pazmiño, D. E., Dudek, H. M. & Fraaije, M. W. Baeyer–Villiger monooxygenases: recent advances and future challenges. *Curr. Opin. Chem. Biol.* **14**, 138–144 (2010).
4. Szolkowy, C., Eltis, L. D., Bruce, N. C. & Grogan, G. Insights into Sequence–Activity Relationships amongst Baeyer–Villiger Monooxygenases as Revealed by the Intragenomic Complement of Enzymes from *Rhodococcus jostii* RHA1. *ChemBioChem* **10**, 1208–1217 (2009).
5. Fraaije, M. W. *et al.* Discovery of a thermostable Baeyer–Villiger monooxygenase by genome mining. *Appl. Microbiol. Biotechnol.* **66**, 393–400 (2004).

6. Reetz, M. T. & Wu, S. Laboratory Evolution of Robust and Enantioselective Baeyer–Villiger Monooxygenases for Asymmetric Catalysis. *J. Am. Chem. Soc.* **131**, 15424–15432 (2009).
7. Bocla, M. *et al.* Converting Phenylacetone Monooxygenase into Phenylcyclohexanone Monooxygenase by Rational Design: Towards Practical Baeyer–Villiger Monooxygenases. *Adv. Synth. Catal.* **347**, 979–986 (2005).
8. Parra, L. P., Acevedo, J. P. & Reetz, M. T. Directed evolution of phenylacetone monooxygenase as an active catalyst for the baeyer–villiger conversion of cyclohexanone to caprolactone. *Biotechnol. Bioeng.* **112**, 1354–1364 (2015).
9. Wu, S., Acevedo, J. P. & Reetz, M. T. Induced allostery in the directed evolution of an enantioselective Baeyer–Villiger monooxygenase. *Proc. Natl. Acad. Sci.* **107**, 2775–2780 (2010).
10. Reetz, M. T. & Wu, S. Greatly reduced amino acid alphabets in directed evolution: making the right choice for saturation mutagenesis at homologous enzyme positions. *Chem. Commun.* 5499 (2008). doi:10.1039/b813388c
11. Sheng, D., Ballou, D. P. & Massey, V. Mechanistic Studies of Cyclohexanone Monooxygenase: Chemical Properties of Intermediates Involved in Catalysis. *Biochemistry (Mosc.)* **40**, 11156–11167 (2001).
12. Pazmiño, D. E. T., Snajdrova, R., Rial, D. V., Mihovilovic, M. D. & Fraaije, M. W. Altering the Substrate Specificity and Enantioselectivity of Phenylacetone Monooxygenase by Structure-Inspired Enzyme Redesign. *Adv. Synth. Catal.* **349**, 1361–1368 (2007).
13. Dudek, H. M. *et al.* Extending the substrate scope of a Baeyer–Villiger monooxygenase by multiple-site mutagenesis. *Appl. Microbiol. Biotechnol.* **98**, 4009–4020 (2014).

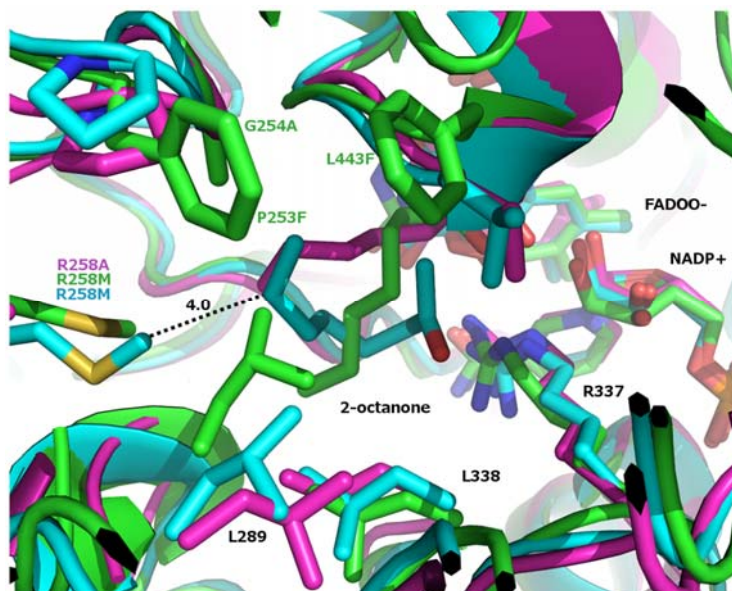
14. Polyak, I., Reetz, M. T. & Thiel, W. Quantum Mechanical/Molecular Mechanical Study on the Mechanism of the Enzymatic Baeyer–Villiger Reaction. *J. Am. Chem. Soc.* **134**, 2732–2741 (2012).
15. Yachnin, B. J., Sprules, T., McEvoy, M. B., Lau, P. C. K. & Berghuis, A. M. The Substrate-Bound Crystal Structure of a Baeyer–Villiger Monooxygenase Exhibits a Criegee-like Conformation. *J. Am. Chem. Soc.* **134**, 7788–7795 (2012).
16. Seo, J.-H. *et al.* Engineering of Baeyer-Villiger monooxygenase-based Escherichia coli biocatalyst for large scale biotransformation of ricinoleic acid into (Z)-11-(heptanoyloxy)undec-9-enoic acid. *Sci. Rep.* **6**, 28223 (2016).
17. Carvalho, A. *et al.* Catalytic mechanism of Phenylacetone monooxygenases for non-native linear substrates: implications on rational engineering of BVMOs to expand the substrate specificity. *Phys. Chem. Chem. Phys.* (2017). doi:10.1039/C7CP03640J
18. Orru, R. *et al.* Snapshots of Enzymatic Baeyer-Villiger Catalysis OXYGEN ACTIVATION AND INTERMEDIATE STABILIZATION. *J. Biol. Chem.* **286**, 29284–29291 (2011).
19. Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).
20. Morris, G. M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).
21. Case, D. *et al.* {Amber 14}. (2014).
22. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).

23. Wang, J. *et al.* Development and testing of a general amber force field. *J Comput Chem* 25–1157 (2004).
24. Iglesias-Arteaga, Martín A., M. A., Velázquez-Huerta, G. A., Méndez-Stivalet, J. M., Galano, A. & Alvarez-Idaboy, J. R. The Baeyer-Villiger reaction of 23-oxosapogenins. *ARKIVOC* **2005**, 109–126
25. Cádenas, R., Reyes, L., Lagúnez-Otero, J. & Cetina, R. Semiempirical studies on the transition structure of the Baeyer and Villiger rearrangement. The reaction of acetone with alkyl and aryl peracids. *J. Mol. Struct. THEOCHEM* **497**, 211–225 (2000).
26. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
27. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models - Miyamoto - 1992 - Journal of Computational Chemistry - Wiley Online Library. Available at: <http://onlinelibrary.wiley.com/doi/10.1002/jcc.540130805/abstract>. (Accessed: 2nd October 2017)
28. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
29. Xia, Y., Chu, W., Qi, Q. & Xun, L. New insights into the QuikChange<sup>TM</sup> process guide the use of Phusion DNA polymerase for site-directed mutagenesis. *Nucleic Acids Res.* gku1189 (2014). doi:10.1093/nar/gku1189
30. Copeland, R. A. *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*. (Wiley-VCH, 2000).

31. Torres Pazmiño, D. E., Baas, B.-J., Janssen, D. B. & Fraaije, M. W. Kinetic Mechanism of Phenylacetone Monooxygenase from *Thermobifida fusca*†. *Biochemistry (Mosc.)* **47**, 4082–4093 (2008).
32. Ferroni, F.M., Smit, M.S., Opperman, D.J. Functional Divergence between closely related Baeyer-Villiger monooxygenases from *Aspergillus flavus*. *Journal of Molecular Catalysis B*, 2014, **107**, 47 - 54.
33. Ferroni, F.M., Tolmie, C., Smit, M.S., Opperman, D.J. Structural and Catalytic Characterization of a Fungal Baeyer-Villiger Monooxygenase. *PLoS One*. 2016;**11**:e0160186.
34. Riebel, A., Dudek, H.M., de, Gonzalo. G., Stepniak, P., Rychlewski, L., Fraaije, M.W. Expanding the set of rhodococcal Baeyer-Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Appl Microbiol Biotechnol*. 2012;**95**(6):1479-89.



## TOC



Using molecular dynamics simulations and steady-state kinetic analysis, we found that the movement of L289 in an active quadruple mutant is crucial for reshaping the catalytic site to enable the non-native aliphatic substrate to adopt a catalytic capable pose in PAMO. The findings would shed light on rational engineering of PAMO to expand its substrate specificity.